

## CHAPTER 2

### MATERIALS & METHODS

#### 2.1 Field work and specimen collections

##### 2.1.1 Sampling strategy and sampling sites

In order to address the phylogenetic relationships within the genus *Alpinia*, the sampling criteria were to obtain as many species as possible to represent the different levels of biogeographic diversity of *Alpinia* species from various ecological regions of Peninsular Malaysia, and to find additional species or taxa lacking in existing clades of previous studies.

Generally, three sampling strategies were used. Firstly, samples were collected from known recorded herbarium localities recorded in selected herbarium departments (Rimba Ilmu Universiti Malaya, FRIM Kepong and Singapore Botanic Gardens). This strategy was used since it has a higher possibility of obtaining the samples in order to reduce the fieldwork time and to verify whether the samples are still available in the recorded localities.

Most of the fieldwork was carried out during recorded flowering and/or fruiting season. For some species where the fruiting and flowering season was not previously recorded, sampling of plants was done opportunistically during the range of other *Alpinia* flowering month.

During the fieldwork, GPS locations were recorded and photographs of *Alpinia* species were taken. Later, whole plants were tagged, cut into sections and paper pressed to be processed later into voucher specimens. Rhizomes of some samples that can be found in abundance were collected to be replanted in the nursery in Institute of Biological Sciences for preservation purposes.

Fruits and flowers were preserved in pre-mixed solutions (a ratio of 3:1 [70% alcohol: acetic acid] or Copenhagen solution (a mixture of Methylated spirit, glycerol and water in a ratio of 10:1:8). In addition, leaf samples were washed with 70% alcohol and distilled water before being cut and dried in silica gel to preserve the DNA for molecular work. A minimum of three replicates were collected for each species.

All species collection were described and initially identified on site based on Holttum's key (1950). Other species which were unknown will be described on site and later verified by Prof. Dr. Halijah Ibrahim.

## **2.2 Laboratory work**

### **2.2.1 Voucher specimens for herbarium collection**

All sample collection which was described and/or identified on site was verified by Prof. Dr. Halijah Ibrahim. Field collections of paper pressed samples were dried for one week in an oven (55 °C) and mounted for voucher purposes. The voucher specimens were either deposited at KLU Herbarium (Rimba Ilmu Universiti Malaya) or FRIM herbarium or Suriana Gardens herbarium .

### **2.2.2 DNA extraction**

DNA was extracted using a modified method of Doyle and Doyle (1987) CTAB (hexadexyltrimethylammonium bromide) protocol or DNeasy Plant Mini Kit (Qiagen, USA) extraction according to manufacturer's instructions. Briefly for CTAB method, preserved leaf samples were grounded using mortar and pestle while continuously adding liquid nitrogen until it formed into powder. The powdered sample was incubated in DNA extraction buffer (CTAB-PVP) and RNase at 65 °C for 45 minutes, cooled to room temperature before being added with chloroform: isoamyl alcohol (in a ratio of

24:1) for 15 minutes prior to centrifuging at 13,000 rpm for 2 minutes. The step of adding chloroform: isoamyl alcohol until centrifugation was repeated twice.

Subsequently, the recovered supernatant was precipitated with ice-cold isopropanol, the DNA obtained washed in 70 % alcohol with 10 mM ammonium acetate. The pellets were vigorously agitated and after 30 minutes were centrifuged using Ebu 12r Hettich Zentrifugen, at 13,000 rpm for 2 minutes and resuspended in Tris-EDTA buffer (pH 8.0). The genomic DNA was then stored at -20 °C. (Please refer to Appendix 1.0 section A and B for general solutions and recipes of DNA extraction solutions and materials).

The extracted genomic DNA was visualized on 1 % (w/v) TBE agarose gel. The purity and concentration of the plant genomic DNA was determined by taking the Optical Density Reading at both  $A_{260}$  and  $A_{280}$  wavelengths using Eppendorf Biophotometer. Genomic DNA with  $A_{260}$  and  $A_{280}$  ratio of 1.6-1.8 were deemed pure and suitable for PCR.

Genomic DNA of different concentrations was diluted to 100 ng prior to PCR. (Please refer to Appendix 1.0 section A and E for recipes of general solutions and recipes of agarose gel and electrophoresis solutions used in this section).

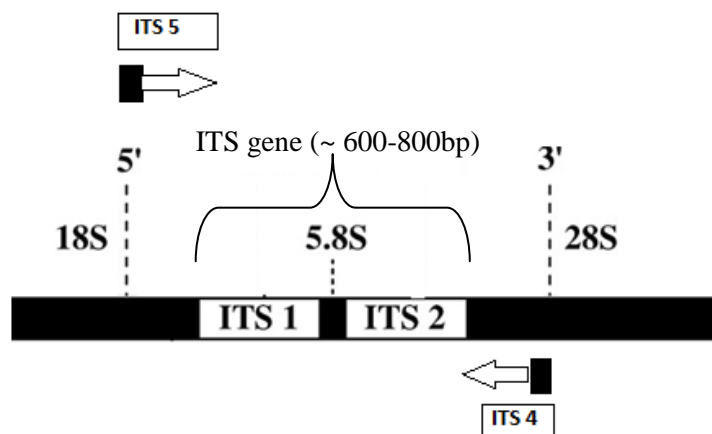
### **2.2.3 PCR Target Regions of the Internal Transcribed Spacer (ITS) region and *maturase K (matK)* chloroplast gene**

In order to compare previous *Alpinia* sequences along with new collected data from this study, two target gene regions (ITS and *matK*) were selected for amplification and sequencing (Figure 2.1 and Figure 2.2). The target regions were initially PCR-amplified to produce template DNA suitable for dye termination sequencing performed

by 1<sup>st</sup> BASE Sequencing Services (Selangor, Malaysia) using Applied Biosystems BigDye® Terminator v3.1 Cycle Sequencing Kit.

#### 2.2.4 PCR amplification of nuclear ITS region

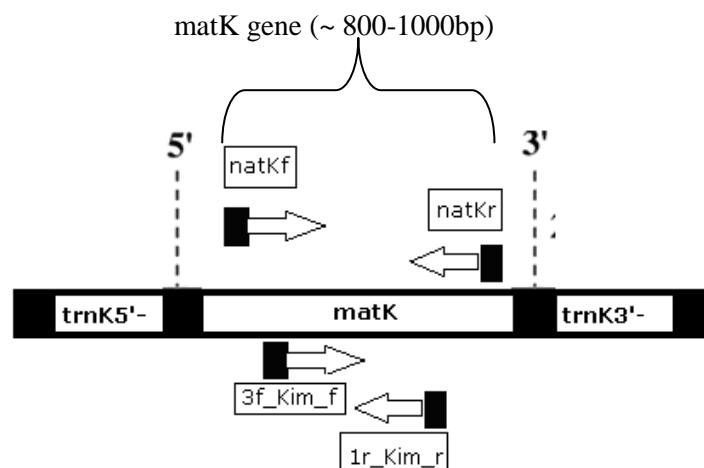
PCR amplification was performed in a 20 µl reaction comprising of 5 µM forward and reverse primers, 2.5 mM each dNTP, 1X i-Taq<sup>TM</sup> DNA Polymerase Reaction buffer, 1U i-Taq<sup>TM</sup> DNA polymerase, 100 ng genomic DNA template and sterilized distilled water up to 20 µl. PCR for the ITS1-5.8S-ITS2 region was performed in an Eppendorf Mastercycler gradient with an initial denaturation of 94 °C for 2 minutes followed by 30 cycles of 94 °C for 30 s, 50-55 °C for 30 s, and 72 °C extension for 45 s, final extension at 72 °C for 2 minutes. PCR products were visualized on 2% (w/v) TBE agarose gel. (Please refer to Appendix 1.0 section C, D and E for oligonucleotide primers, PCR reagents used for PCR amplification and recipes of agarose gels and electrophoresis solutions).



**Figure 2.1 Amplification and sequencing strategy of the nuclear ITS gene.** Bars indicate the annealing positions of primers used (in boxes) whereas the arrows show the directions of the sequence fragments obtained.

## 2.2.5 PCR amplification of partial *matK* region

PCR amplification of partial *matK* gene of the middle section was performed initially with newly designed primer set 1 (natKf and natKr) because the final product is more than 1kb; unless if it fails to amplify, another set of primer were used (primer set 2: 3f\_Kim\_f and 1r\_Kim\_r) which yield a shorter sequence of about 800bp. Each tube of PCR was done in a 20  $\mu$ l reaction comprising of 1  $\mu$ M forward and reverse primers, 0.2 mM final dNTPs, 1X i-Taq<sup>TM</sup> DNA Polymerase Reaction buffer, 1.5 mM MgCl<sub>2</sub>, 2U i-Taq<sup>TM</sup> DNA polymerase, 4% DMSO of total volume, 0.1 mg/ml BSA, 100 ng genomic DNA template and sterilized distilled water up to 20  $\mu$ l. PCR for the *matK* chloroplast region was performed in an Eppendorf Mastercycler gradient with an initial denaturation of 95 °C for 2.5 minutes followed by 10 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s. Then, 25 cycles of 88 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s and final 72 °C extension for 10 minutes. PCR products were visualized on 2% (w/v) TBE agarose gel. (Please refer to Appendix 1.0 section C, D and E for oligonucleotide primers, PCR reagents used for PCR amplification and recipes of agarose gels and electrophoresis solutions).



**Figure 2.2 Amplification and sequencing strategy of the chloroplast *matK* gene.** Bars indicate the annealing positions of primers used (in boxes) whereas the arrows show the directions of the sequence fragments obtained.

## **2.2.6 Purification of PCR products and DNA sequencing**

PCR products or DNA fragments corresponding to the expected size for ITS and *matK* regions were isolated from agarose gel and purified using Megaquickspin™ PCR & Agarose Gel DNA extraction system (iNtRON Biotechnology, Korea) according to the manufacturer's instructions. The purified DNA fragments were visualized on 2 % (w/v) TBE agarose gel. (Please refer to Appendix 2.0 section E for recipes of agarose gels and electrophoresis solutions).

The results of the PCR products were then sent for sequencing. DNA sequencing for ITS and *matK* region was performed by 1<sup>st</sup> BASE Sequencing Services (Selangor, Malaysia) using Applied Biosystems BigDye® Terminator v3.1 Cycle Sequencing Kit. Forward and reverse primers were used as sequencing primers to create strong and reliable consensus sequence.

## **2.3 Molecular Sequence Analysis**

### **2.3.1 Construction of contig sequences**

ITS and *matK* sequences obtained from 1<sup>st</sup> BASE Sequencing Services (Selangor, Malaysia) were manually checked and edited accordingly. Later, contig sequences for each taxa were constructed using the respective forward and reverse sequences and submitted to BLAST (Altschul *et al.*, 1990) using MEGABLAST option to verify the putative species and also to evade sequences from contaminated samples.

### **2.3.2 Alignment of DNA sequences**

Sequences of taxa were aligned separately for each gene using ClustalW in MEGA 5 (Tamura *et al.*, 2011). Alignments were inspected and confirmed by eye and

any ambiguous alignments were corrected manually. Primer sequences flanking the region of interest were removed to obtain a uniformed alignment. The sequence alignment was trimmed at both ends. The aligned ITS and *matK* sequences of the *Alpinia* species of Peninsular Malaysia were submitted to Genbank database in batches using Sequin 10.3 (National Centre for Biotechnology Information, 2010).

The consensus sequences for aligned ITS and *matK* (.mas) were exported in FASTA format (.fas) to analyse its sequence variation. The alignment files were also exported into Nexus format (.nex) for analysis in PAUP4.0b10 (Swofford, 2002).

Sequence variation, nucleotide composition, substitutions rate and genetic distances were computed using application in MEGA 5 (Tamura *et al.*, 2011), whereas the distribution of variable sites in ITS and *matK* sequences were plotted using sliding window application in DnaSP (Librado and Rozas, 2009).

## **2.4 Phylogenetic Analysis**

For phylogenetic analysis, additional sequences for ingroup and outgroup were obtained from GenBank for both genes (Table 2.1). Alignment of sequences were done separately for both genes and also trimmed at both ends to achieve maximum homology in the final sequence alignments. Please refer to Appendix 3.0 section A & B for list of reference sequences used for ITS and *matK* phylogenetic trees.

Phylogenetic trees were constructed separately for ITS and *matK* using neighbour joining analysis, maximum parsimony and Bayesian analysis. Prior to building the concatenated tree using the combined sequences of ITS and *matK*, incongruence length difference (ILD) test was performed. Using the partition homogeneity function in PAUP\*4.0 (Swofford, 2002), the result gave a p-value of 0.01,

indicating incongruence between the gene sequences. Therefore, the phylogenetic trees of combined genes were not constructed due to high differences between both genes.

#### **2.4.1 Choosing the evolutionary model**

Likelihood scores of 56 models were computed initially using PAUP\*4.0 (Swofford, 2002) and the scores was incorporated in Modeltest3.7 (Posada & Crandall, 1998), to search for the evolutionary model of ITS and *matK* sequences based on the Akaike Information Criteria (AIC). The evolutionary models were chosen using jModelTest 2.0 (Darriba *et al.*, 2012) whereby likelihood scores of 88 models were directly computed from the software.

#### **2.4.2 Construction of Neighbour Joining (NJ) tree**

NJ analysis was implemented in PAUP 4.0\* (Swofford, 2002) for Macintosh. The corresponding evolutionary model obtained from jModeltest 2.0 (Darriba *et al.*, 2012) was appended in the Nexus file of ITS and *matK* respectively. Outgroup was defined and made into a monophyletic sister group to the ingroup. Analysis was set to Distance analysis and Distance settings following the Maximum Likelihood Distance as indicated in the evolutionary model. Bootstrap searches were conducted with 1000 replicates and resampling of characters. Majority rule consensus tree was generated with a cut-off value set at 70% bootstrap support.



### **2.4.3 Construction of Maximum Parsimony (MP) tree**

MP analysis was implemented in PAUP 4.0\* (Swofford, 2002) for Macintosh. The corresponding evolutionary model obtained from jModeltest 2.0 (Darriba *et al.*, 2012) was appended in the Nexus file of ITS and *matK* respectively. Outgroup was defined and made into a monophyletic sister group to the ingroup. Analysis was set to Parsimony analysis and a tree-bisection-reconnection (TBR) branch swapping algorithm, STEEPEST DESCENT off, MULTREES on, COLLAPSE branches if maximum length is zero (AMB), with random sequence addition of 10 replicates was used for the MP heuristic search. Bootstrap searches were conducted with 1000 replicates and resampling of characters. Majority rule consensus tree was generated with a cut-off value set at 70% bootstrap support.

### **2.4.4 Construction of Bayesian tree**

BI analysis was implemented in Mr. Bayes 3.2.2 (Ronquist and Huelsenbeck, 2003). The corresponding evolutionary model including the rate matrix and base frequencies obtained from jModeltest 2.0 (Darriba *et al.*, 2012) were appended in the Nexus file of ITS and *matK* respectively. The analysis was run using four Markov Chain Monte Carlo (MCMC) chains with tree samplings of every 100th generations. The first 25% of the generated trees were discarded, and the remainder were used to create a majority rule consensus tree where branches with more than 0.70 posterior probabilities were retained.

#### **2.4.5 Sequence divergence**

DNA sequence divergences of intra- and inter-species were calculated for ITS and *matK* alignment in MEGA5.0 (Tamura *et al.*, 2011). The average and the range of sequence divergences for each recognised clade were calculated in Microsoft Excel using the inter-species divergences.

**Table 2.1. List of species and corresponding Genbank accession numbers for internal transcribed spacer region (ITS) and maturase K (*matK*) gene sequences for construction of phylogenetic trees.** Reference sequences from Genbank (a – Kress *et al.*, 2002; b – Kress *et al.*, 2005; c - Li *et al.*, 2011; d - Rangsiroji *et al.*, 2000)

Species	Genbank accession number	
	ITS	<i>matK</i>
<b>Ingroup</b>		
<i>Alpinia abundiflora</i>	AY742334.1 <sup>b</sup>	AY742393.1 <sup>b</sup>
<i>Alpinia aquatica</i>	AY742335.1 <sup>b</sup>	JF953079.1 <sup>c</sup>
<i>Alpinia arctiflora</i>	AY742336 <sup>b</sup>	AY742395.1 <sup>b</sup>
<i>Alpinia boia</i>	AY742340.1 <sup>b</sup>	–
<i>Alpinia calcarata</i>	AF478710 <sup>a</sup>	AF478810.1 <sup>a</sup>
<i>Alpinia conchigera</i>	AF478712.1 <sup>a</sup>	AF478812.1 <sup>a</sup>
<i>Alpinia eremochlamys</i>	AY742346.1 <sup>b</sup>	AY742404.1 <sup>b</sup>
<i>Alpinia fax</i>	AY742348.1 <sup>b</sup>	AY742405.1 <sup>b</sup>
<i>Alpinia galanga</i>	AF478715.1 <sup>a</sup>	AF478815.1 <sup>a</sup>
<i>Alpinia glabra</i>	AB097221 <sup>b</sup>	JF715463.1 <sup>b</sup>
<i>Alpinia intermedia</i>	AF478716 <sup>a</sup>	AF478816.1 <sup>a</sup>
<i>Alpinia japonica</i>	AY742357 <sup>b</sup>	AY742412.1 <sup>b</sup>
<i>Alpinia javanica</i>	AY742358 <sup>b</sup>	AY742413.1 <sup>b</sup>
<i>Alpinia latilabris</i>	AY742360 <sup>b</sup>	AY742414.1 <sup>b</sup>
<i>Alpinia ligulata</i>	AY742361 <sup>b</sup>	AY742415.1 <sup>b</sup>
<i>Alpinia malaccensis</i>	AH008309 <sup>d</sup>	AY742417.1 <sup>d</sup>
<i>Alpinia modesta</i>	AY742364 <sup>b</sup>	AY742418.1 <sup>b</sup>
<i>Alpinia mutica</i>	AY742365 <sup>b</sup>	AY742420.1 <sup>b</sup>
<i>Alpinia nutans</i>	AY742369 <sup>b</sup>	AY742417.1 <sup>b</sup>
<i>Alpinia officinarum</i>	AF478718 <sup>a</sup>	AF478818.1 <sup>a</sup>
<i>Alpinia oxymitra</i>	AY742371 <sup>b</sup>	–
<i>Alpinia oxyphylla</i>	AY742372 <sup>b</sup>	AF478818.1 <sup>b</sup>
<i>Alpinia pinetorum</i>	AY742373 <sup>b</sup>	AF478820.1 <sup>b</sup>
<i>Alpinia polyantha</i>	AY745692 <sup>b</sup>	AF478820.1 <sup>b</sup>
<i>Alpinia rafflesiana</i>	AY742376 <sup>b</sup>	AY742430.1 <sup>b</sup>
<i>Alpinia rubricaulis</i>	AY742378 <sup>b</sup>	AY742432.1 <sup>b</sup>
<i>Alpinia sibuyanensis</i>	AY742381.1 <sup>b</sup>	AY742431.1 <sup>b</sup>
<i>Alpinia vittata</i>	AF478720.1 <sup>a</sup>	AF478820.1 <sup>a</sup>
<i>Alpinia zerumbet</i>	AY742389.1 <sup>b</sup>	AY742443.1 <sup>b</sup>
<b>Outgroup</b>		
<i>Burbridgea nitida</i>	AF478728.1 <sup>a</sup>	AF478828.1 <sup>a</sup>
<i>Pleuranthodium floccosum</i>	AY742333.1 <sup>a</sup>	AY742391.1 <sup>a</sup>
<i>Riedelia</i> sp.	AF478785.1 <sup>a</sup>	AF478886.1 <sup>a</sup>